D1 al. (1995) Am. J. Hum. Genet. *57*, 1384-1394) identified two regions on chromosome 18 showing linkage to the disease.

Please replace the fourth paragraph on page 1 with:

The present invention is based on our discovery of a novel gene which maps to 18q21 and which unexpectedly shows appreciable sequence homology to the ned-4 gene on chromosome 15. Ned-4 is the human homologue of the mouse nedd-4 gene which is known to be differentially expressed during neural development and to be involved in signal transduction. Human ned-4 has been shown (Schild L., Lu Y., Gautshi I., Schneeberger E., Lifton R.P. and Rossier B.C. (1996) EMBO J. *15*, 2381-2387; Straub O., Dho S., Henry P.C., Correa J., Ishikawa T., McGlade J. and Rotin D. (1996) EMBO J *15*, 2371-2380) to be a negative regulator of a sodium channel which is deleted in Liddle's syndrome (a hereditary form of hypertension).

D2

Please replace the first complete paragraph on page 2 with:

Nedd-4 was originally isolated as a partial cDNA clone from a mouse brain library (Kumar S., Tomooka Y. and Noda M. (1992) Biochem. Biophys. Res. Commun. *185*, 1155-1161) as one of a set of genes which were differentially expressed during development (Neural precursor cells expressed developmentally down-regulated). The derived amino acid sequence contains three copies of the WW domain (Andre B. and Springael J.Y. (1994) Biochem. Biophys. Res. Commun. *205*, 1201-1205; Bork P. and Sudol M. (1994) Trends Biochem.Sci. *19*, 531-533; Hofmann K. and Boucher P. (1995) FEBS Letts. *358*, 153-157), a Ca lipid binding (CaLB/C2) domain (Brose N., Hofmann K.O., Hata Y., Suedhof T.C. (1995) J. Biol. Chem. *270*, 25273-25280) and a Hect (homologous to the E6-AP carbodyl terminus) domain which has homology to a ubiquitin ligase (E3) enzyme (Huibregtse J.M., Scheffner M., Beaudenon S. and Howley P.M. (1995) Proc. Natl. Acad. Sci. USA, *92*, 2563-2567). The human homologue of nedd-4 (Ned-4) was isolated as an randomly cloned EST (KIAA0093) from immature myeloblast mRNA (Nomura N. et al. (1994) DNA Res. *2*, 37-43) and shown by sequence comparison to have 86% identity at the amino acid level to the mouse sequence. The human sequence, however, has a fourth copy of the WW domain.

D3

Please replace the second complete paragraph on page 2 with:

The WW domain is a 40 amino acid sequence found in several unrelated proteins. The two highly conserved tryptophans give it its name. The function of the domain is thought to be involved in protein-protein interactions. Despite their functional diversity, the proteins listed all appear to be involved in cell signalling or regulation. It has been shown that the WW domains of Nedd-4 interact with the proline-rich PY motifs in the epithelial sodium channel in the kidney (Schild L., Lu Y., Gautschi I., Schneeberger E., Lifton R.P. and Rossier B.C. (1996) EMBO J. *15*, 2381-2387). Mutational deletion of the PY motifs in the epithelium sodium channel in Liddle's syndrome, an inherited disease causing systemic hypertension characterised by hyperactivity of the sodium channel, has been shown to abrogate binding of Nedd-4 (Straub O., Dho S., Henry P.C., Correa J., Ishikawa T., McGlade J. and Rotin D. (1996) EMBO J. *15*, 2371-2380). It is therefore likely that Nedd-4 has a negative regulatory role when bound to the channel.

◊◊ Please replace the second complete paragraph on page 4 with:

In a further aspect of the invention we provide a recombinant ZGGBP1 protein obtained by expression of all or a part of the cDNA as set out in SEQ ID NO: 1. The recombinant protein may comprise all or a convenient part of the peptide sequence set out in SEQ ID NO: 2. The production of a protein according to the invention may be achieved using standard recombinant DNA techniques involving the expression of the protein by a host cell as described for example by Sambrook J., Fritsch E.F. and Maniatis T. (1989) "Molecular Cloning: A Laboratory Manual (2nd edition)" Cold Spring Harbor Laboratory Press NY. The isolated nucleic acids described herein may for example be introduced into any convenient expression vector for example the T7 Studier system for expression in E.coli (US-A-4952496), Pichia pastoris for expression in yeast, the Baculovirus system for expression in insect cells and the GS system for expression in mammalian cells by operatively linking the DNA to any necessary expression control elements therein and transforming any suitable prokaryotic or eukaryotic host cell with the vector using well known procedures.

◊◊ Please replace the description of the figures from page 7, line 29, to page 8, line 14, with:

Figure 1 shows the predicted amino acid sequence (SEQ ID NO:2) of ZGGBP1. The C2 domain is indicated by carets, the four WW domains are indicated by asterisks and the Hect domain is indicated by underlining.

Figure 2 shows a comparison of amino acid sequences of human ned4 Swissprot entry P46934 (SEQ ID NO: 6) and ZGGBP1 (SEQ ID NO: 2).

Figure 3 shows a Northern blot analysis of various human tissues probed with ZGGBP1.

Figure 4 shows a comparison of the nucleic acid sequences of human (SEQ ID NO: 7) and mouse (SEQ ID NO:3) ZZGBP1. The mouse sequence is a partial cDNA which spans the C-terminal portion of the human protein coding region.

Figure 5 shows a comparison of the nucleic acid sequences for ZGGBP1 (SEQ ID NO: 1) and Pub 3 (SEQ ID NO: 8).

Figure 6 shows a polymorphism located at position 3554 of the cDNA sequence (SEQ ID NO: 10) (SEQ ID NO: 11).

Figure 7 shows a polymorphism located at position 4828 of the cDNA sequence (SEQ ID NO: 12) (SEQ ID NO: 13).

Figure 8 shows a polymorphism located in an intronic sequence (SEQ ID NO:14) (SEQ ID NO: 15) derived from a BAC containing ZGGBP1.

Figure 9 shows a variable number of tetranucleotide repeats located within an intronic sequence (SEQ ID NO:9) from ZGGBP1.

Figure 10 shows an insertion at position 4032 of the cDNA sequence (SEQ ID NO: 16) (SEQ ID NO:17).

Please delete the reference list from page 8, line 16, to page 9, line 9.

Please replace the second complete paragraph on page 9 with:

The 18q21 region described by Stine O.C. et al. (1994) Am. J. Hum. Genet. 57, 1384-1394 is delimited by the STS markers used by that group to identify linkage. They found the most strongly linked marker to be D18S41, which had a LOD score of 3.51 in cases of paternal inheritance. Linkage declined over flanking markers. We identified a set of four Yeast Artificial Chromosomes (YACs) which comprised a contiguous overlapping set of genomic clones covering the defined region by the presence in those YACs of STS markers used in the Stine study.

◊◊ Please replace the third complete paragraph on page 9 with:

DNA from the YACs was prepared and used in a PCR-based hybridisation approach to entrich for transcripts from a human fetal brain DNA library. This approach, known as direct selection (Lovett M., Kere J. and Hinton L. (1991) Proc. Natl. Acad. Sci. USA *88*, 9628-9632) has been shown to be efficient in identifying transcripts present on large genomic clones.

◊◊ Please replace the first complete paragraph on page 10 with:

The UNIGENE database is a repository for transcripts which have been mapped by taking representative Expressed Sequence Tagged Sites (ESTs) and performing PCR analysis on a parcel of radiation hybrids which have been calibrated with respect to a framework of 1000 genetic markers (Schuler G.D. et al. (1996) Science *274*, 540-546). We found 36 EST clusters which had been mapped to a radiation hybrid map interval which corresponded to the 18q21 region of interest and to flanking regions outside.

◊◊ Please replace the first complete paragraph on page 11 with:

The extending of partial transcripts to full length clones can be a complex and difficult process requiring skill and expertise for success. Having considered several possibilities, we opted for a PCR-based approach to isolate and characterise the full length ZGGBP1 gene. Human foetal brain double stranded cDNA was synthesised from mRNA using standard methods (Sambrook et al. 1989) and ligated into lambda Zap vector by use of adapters. However, in order to minimise the loss of transcripts often seen following the cloning step, the resulting ligation mix was not cloned but was instead used as a template for PCR. Oligonucleotide primers specific to ZGGBP1 were used in combination with vector specific primers to amplify

DNA across the unknown part of the gene. Since the distance to be covered was unknown, we performed long PCR using the commercially available BCL Expand enzyme and long (30mer) oligonucleotide primers. Since we were using unamplified material, where our target cDNAs were likely to be present only in very small amounts, we utilised a secondary PCR step with nested oligonucleotide primers and again using long PCR to yield sufficient PCR products to be visible by gel analysis and also to minimise the possibility of non-specific PCR amplification. The PCR products derived from these experiments were then purified and sequenced directly. Where necessary, the DNA sequence obtained was used to design further primers to walk along the gene in a 3' - 5' direction. The complete nucleotide sequence derived from this work is 5.2kb and the translated amino acid sequence is shown in SEQ ID NO: 2.

◊◊ Please replace the third complete paragraph on page 12 with:

The Research Genetics human Bacterial Artificial Chromosome (BAC) library (Shizua H., Birren B., Kim U-J., Mancino V., Slepak T., Tachiiri Y. and Simon M. (1992) Proc. Natl. Acad. Sci. USA *89*, 8794-8797; Kim U-J., Shizuya H., Kang H-L., Choi S-S., Garrett C.L., Smink L.J., Birren B.W., Korenberg J.R., Dunham I. and Simon M. (1996) Proc. Natl. Acad. Sci. USA *93*, 6297-6301) was screened by PCR using primers specific to the 3'UTR of ZGGBP1 and BACs were isolated. These are being used to characterise the structural gene including the intronexon structure and the 5' regulatory region.

A marked-up version of the amended paragraphs is provided below.

Family, twin and adoption studies have suggested the importance of genetic predisposition to bipolar affective disorder. On this basis, several groups have undertaken genetic linkage analysis in families with a high incidence of the disorder to find a causal gene. Many of the studies show conflicting data suggesting that a single gene is unlikely to be the cause. Rather, multiple interacting genetic traits may be involved. A recent study (~~Stine et al. 1995~~ Stine O.C. et al. (1995) Am. J. Hum. Genet. *57*, 1384-1394) identified two regions on chromosome 18 showing linkage to the disease.

The present invention is based on our discovery of a novel gene which maps to 18q21 and which unexpectedly shows appreciable sequence homology to the ned-4 gene on chromosome 15. Ned-4 is the human homologue of the mouse nedd-4 gene which is known to be differentially expressed during neural development and to be involved in signal transduction. Human ned-4 has been shown (~~Schild et al. 1996, Straub et al. 1996~~ Schild L., Lu Y., Gautshi I., Schneeberger E., Lifton R.P. and Rossier B.C. (1996) EMBO J. *15*, 2381-2387; Straub O., Dho S., Henry P.C., Correa J., Ishikawa T., McGlade J. and Rotin D. (1996) EMBO J *15*, 2371-2380) to be a negative regulator of a sodium channel which is deleted in Liddle's syndrome (a hereditary form of hypertension).

Nedd-4 was originally isolated as a partial cDNA clone from a mouse brain library (~~Kumar et al. 1992~~Kumar S., Tomooka Y. and Noda M. (1992) Biochem. Biophys. Res. Commun. *185*, 1155-1161) as one of a set of genes which were differentially expressed during development (Neural precursor cells expressed developmentally down-regulated). The derived amino acid sequence contains three copies of the WW domain (~~Andre & Springael 1994, Bork & Sudol,1994; Hofman & Boucher, 1995~~Andre B. and Springael J.Y. (1994) Biochem. Biophys. Res. Commun. *205*, 1201-1205; Bork P. and Sudol M. (1994) Trends Biochem.Sci. *19*, 531-533; Hofmann K. and Boucher P. (1995) FEBS Letts. *358*, 153-157), a Ca lipid binding (CaLB/C2) domain (Brose N., Hofmann K.O., Hata Y., Suedhof T.C. (1995) J. Biol. Chem. *270*, 25273-25280) and a Hect (homologous to the E6-AP carbodyl terminus) domain which has homology to a ubiquitin ligase (E3) enzyme (~~Huibregtse et al. 1995~~ Huibregtse J.M., Scheffner M., Beaudenon S. and Howley P.M. (1995) Proc. Natl. Acad. Sci. USA, *92*, 2563-2567). The human homologue of nedd-4 (Ned-4) was isolated as an randomly cloned EST (KIAA0093) from immature myeloblast mRNA (~~Nomura N. et al. 1994~~ Nomura N. et al. (1994) DNA Res. *2*, 37-43) and shown by sequence comparison to have 86% identity at the amino acid level to the mouse sequence. The human sequence, however, has a fourth copy of the WW domain.

The WW domain is a 40 amino acid sequence found in several unrelated proteins. The two highly conserved tryptophans give it its name. The function of the domain is thought to be

involved in protein-protein interactions. Despite their functional diversity, the proteins listed all appear to be involved in cell signalling or regulation. It has been shown that the WW domains of Nedd-4 interact with the proline-rich PY motifs in the epithelial sodium channel in the kidney (~~Schild et al. 1996~~ Schild L., Lu Y., Gautschi I., Schneeberger E., Lifton R.P. and Rossier B.C. (1996) EMBO J. *15*, 2381-2387). Mutational deletion of the PY motifs in the epithelium sodium channel in Liddle's syndrome, an inherited disease causing systemic hypertension characterised by hyperactivity of the sodium channel, has been shown to abrogate binding of Nedd-4 (~~Straub et al. 1996~~ Straub O., Dho S., Henry P.C., Correa J., Ishikawa T., McGlade J. and Rotin D. (1996) EMBO J. *15*, 2371-2380). It is therefore likely that Nedd-4 has a negative regulatory role when bound to the channel.

In a further aspect of the invention we provide a recombinant ZGGBP1 protein obtained by expression of all or a part of the cDNA as set out in SEQ ID NO: 1. The recombinant protein may comprise all or a convenient part of the peptide sequence set out in SEQ ID NO: 2. The production of a protein according to the invention may be achieved using standard recombinant DNA techniques involving the expression of the protein by a host cell as described for example by ~~Sambrook et al. 1989~~ Sambrook J., Fritsch E.F. and Maniatis T. (1989) "Molecular Cloning: A Laboratory Manual (2nd edition)" Cold Spring Harbor Laboratory Press NY. The isolated nucleic acids described herein may for example be introduced into any convenient expression vector for example the T7 Studier system for expression in E.coli (US-A-4952496), Pichia pastoris for expression in yeast, the Baculovirus system for expression in insect cells and the GS system for expression in mammalian cells by operatively linking the DNA to any necessary expression control elements therein and transforming any suitable prokaryotic or eukaryotic host cell with the vector using well known procedures.

Figure 1 shows the predicted amino acid sequence (SEQ ID NO:2) of ZGGBP1. The C2 domain is indicated by carets, the four WW domains are indicated by asterisks and the Hect domain is indicated by underlining.

Figure 2 shows a comparison of amino acid sequences of human ned4 Swissprot entry P46934 (SEQ ID NO: 6) and ZGGBP1 (SEQ ID NO: 2).

Figure 3 shows a Northern blot analysis of various human tissues probed with ZGGBP1.

Figure 4 shows a comparison of the nucleic acid sequences of human (SEQ ID NO: 7) and mouse (SEQ ID NO:3) ZZGBP1. The mouse sequence is a partial cDNA which spans the C-terminal portion of the human protein coding region.

Figure 5 shows a comparison of the nucleic acid sequences for ZGGBP1 (SEQ ID NO: 1) and Pub 3 (SEQ ID NO: 8).

Figure 6 shows a polymorphism located at position 3554 of the cDNA sequence (SEQ ID NO: 10) (SEQ ID NO: 11).

Figure 7 shows a polymorphism located at position 4828 of the cDNA sequence (SEQ ID NO: 12) (SEQ ID NO: 13).

Figure 8 shows a polymorphism located in an intronic sequence (SEQ ID NO:14) (SEQ ID NO: 15) derived from a BAC containing ZGGBP1.

Figure 9 shows a variable number of tetranucleotide repeats located within an intronic sequence (SEQ ID NO:9) from ZGGBP1.

Figure 10 shows an insertion at position 4032 of the cDNA sequence (SEQ ID NO: 16) (SEQ ID NO:17).

The 18q21 region described by ~~Stine et al. (1995)~~ Stine O.C. et al. (1994) Am. J. Hum. Genet. 57, 1384-1394 is delimited by the STS markers used by that group to identify linkage. They found the most strongly linked marker to be D18S41, which had a LOD score of 3.51 in cases of paternal inheritance. Linkage declined over flanking markers. We identified a set of four Yeast Artificial Chromosomes (YACs) which comprised a contiguous overlapping set of

genomic clones covering the defined region by the presence in those YACs of STS markers used in the Stine study.

DNA from the YACs was prepared and used in a PCR-based hybridisation approach to entrich for transcripts from a human fetal brain DNA library. This approach, known as direct selection (~~Lovett et al. 1991~~ Lovett M., Kere J. and Hinton L. (1991) Proc. Natl. Acad. Sci. USA *88*, 9628-9632) has been shown to be efficient in identifying transcripts present on large genomic clones.

The UNIGENE database is a repository for transcripts which have been mapped by taking representative Expressed Sequence Tagged Sites (ESTs) and performing PCR analysis on a parcel of radiation hybrids which have been calibrated with respect to a framework of 1000 genetic markers (~~Schuler et al. 1996~~ Schuler G.D. et al. (1996) Science *274*, 540-546). We found 36 EST clusters which had been mapped to a radiation hybrid map interval which corresponded to the l8q21 region of interest and to flanking regions outside.

The extending of partial transcripts to full length clones can be a complex and difficult process requiring skill and expertise for success. Having considered several possibilities, we opted for a PCR-based approach to isolate and characterise the full length ZGGBP1 gene. Human foetal brain double stranded cDNA was synthesised from mRNA using standard methods (Sambrook et al. 1989) and ligated into lambda Zap vector by use of adapters. However, in order to minimise the loss of transcripts often seen following the cloning step, the resulting ligation mix was not cloned but was instead used as a template for PCR. Oligonucleotide primers specific to ZGGBP1 were used in combination with vector specific primers to amplify DNA across the unknown part of the gene. Since the distance to be covered was unknown, we performed long PCR using the commercially available BCL Expand enzyme and long (30mer) oligonucleotide primers. Since we were using unamplified material, where our target cDNAs were likely to be present only in very small amounts, we utilised a secondary PCR step with

nested oligonucleotide primers and again using long PCR to yield sufficient PCR products to be visible by gel analysis and also to minimise the possibility of non-specific PCR amplification. The PCR products derived from these experiments were then purified and sequenced directly. Where necessary, the DNA sequence obtained was used to design further primers to walk along the gene in a 3' - 5' direction. The complete nucleotide sequence derived from this work is 5.2kb and the translated amino acid sequence is shown in SEQ ID NO: ~~1~~ 2.

The Research Genetics human Bacterial Artificial Chromosome (BAC) library (~~Shizua et al. 1992, Kim et al. 1996~~ Shizua H., Birren B., Kim U-J., Mancino V., Slepak T., Tachiiri Y. and Simon M. (1992) Proc. Natl. Acad. Sci. USA *89*, 8794-8797; Kim U-J., Shizuya H., Kang H-L., Choi S-S., Garrett C.L., Smink L.J., Birren B.W., Korenberg J.R., Dunham I. and Simon M. (1996) Proc. Natl. Acad. Sci. USA *93*, 6297-6301) was screened by PCR using primers specific to the 3'UTR of ZGGBP1 and BACs were isolated. These are being used to characterise the structural gene including the intronexon structure and the 5' regulatory region.

In the claims:

For the convenience of the Examiner, all claims being examined, whether or not amended, are presented below.

Please cancel, without prejudice, claim 22.

10. (**Amended**) A purified polypeptide comprising the amino acid sequence of SEQ ID NO:2.

11. (**Amended**) A purified polypeptide comprising the amino acid sequence of SEQ ID NO:4.

12. (**Reiterated**) A fusion protein in which a polypeptide of claim 10 or claim 11 is fused with glutathione-S-transferase.